What Performance Assessments Do and Why We Need Them in Schools

by Anthony Petrosky

What Performance Assessments Do

Performance assessments were originally created to assess people's abilities to do complex activities that are important to their jobs. Surgeons collect electronic portfolios of their surgeries on life-like mannequins or in computer simulations. Architects create plans for buildings in extensive portfolios of drawings. Designers make drawings to specifications for clothes and present them in portfolios to buyers. Camera operators put together reels of movie clips that demonstrate their abilities to frame scenes, follow focus, and track action. Flight control radar operators demonstrate their skills through computer simulations of challenging scenarios with multiple planes flying on various routes at various speeds.

Science writers create sets of essays on complex topics to demonstrate their knowledge and abilities to write for ordinary people. Artists create portfolios of their drawings or photographs to demonstrate the range and depth of their work to others, including gallery owners who might be interested in showing their pieces. Firemen and police officers regularly participate in mock performances that mirror situations in which they often find themselves. Sometimes these are videotaped so that the participants can assess their performances against specific criteria.

Performance Assessments of Students' Work

Performance assessments for students usually ask them to do the sorts of academic work that defines their subjects. In English, they write various sorts of essays and often collect them into portfolios to demonstrate their work over time and with different types of writing tasks. In mathematics, students respond to word problems that represent big mathematical concepts such as ratio and proportion by solving the problems and explaining their solutions in writing. These mathematic performances might also be collected in portfolios to demonstrate work over time since it often takes multiple performances to assess conceptual success.

The emphasis in performance assessments is usually on multiple demonstrations of work samples that can be evaluated or judged, often as they represent the work over periods of time, against criteria for excellence and satisfactory performance. Such criteria, when applied to complex performance and portfolios of work samples, can be diagnostic for teachers in ways that are impossible with multiple choice tests.
Teachers can learn, for instance, that over time and on multiple essays, students struggle with explaining how textual evidence connects to the points they’re making in their essays. There are no multiple choice tests that can give teachers this kind of diagnostic information.

**Performance versus Identification and Regurgitation**

The strength of performance assessments comes from their abilities to capture authentic examples of work samples rather than proxies for work samples. We know, for example, that students' scores on multiple choice tests of editing skills in writing generally correlate with students' writing abilities, but the multiple choice tests are proxies for the students' writing—not authentic assessments of the writing. Many assessment experts think that such multiple choice tests of editing are not even good indications of a students' editing abilities because they are being asked to edit contrived and often tricky sentences not of their own making.

Test makers argue that the proxies are cheaper and good enough for assessing writing; however, tremendous problems of practice emerge from our willingness to accept such proxies for authentic performances. During the No Child Left Behind (NCLB) testing era, students took state-wide, high-stakes, multiple choice tests and very few were actually asked to write authentic essays. When students did write essays, the subjects were frequently either unfamiliar to them or uninspiring. It’s difficult to gauge students’ commitments to performing well on such tests.

**The Result of Multiple Choice Testing**

The end results of all this multiple choice testing in the US has been deeply documented over and over—it’s led to months of test preparation in which students complete exercises for weeks on end that look like the ones they’ll take on the tests. Instruction has bent and shaped itself to prepare students for these multiple choice tests in which the emphasis in class is on identifying and regurgitating information—as it is on the tests—rather than on creating or making summaries or explanations or arguments or even poems in writing or in talk.

The nation turned away from its large scale experiments with performance assessments in writing during the 1990s because of the expense involved in gathering and preparing teachers to rate portfolios against criteria. The National Board for Professional Teaching Standards (NBPTS), to its credit, has maintained its focus on assessing teachers for board certification through sophisticated portfolios of practice. In addition, the National Assessment of Education Practice (NAEP), our nation’s best assessment effort to date, continues to solicit writing samples from students that are scored against relevant criteria.
Lately, and partly in response to the ways that instruction has been driven to look like multiple choice testing, performance assessments are making a comeback. The two major national assessment consortia—PARCC and Smarter Balanced—have incorporated sophisticated exercises in their assessments that ask students to write evidence-based explanations and arguments based on their readings of single and multiple texts on the same topics. Some states, such as Virginia, have begun to explore alternative approaches to these top-down assessments by creating local mandates that empower teachers to create curriculum-embedded performance assessments and portfolios of writing that assess the concepts and skills in the curricula that are being taught.

The Big Take Away

The big take away from our history with multiple choice and performance assessments is that we know assessment drives instruction to look like the assessments. It’s no wonder that students tested over and over on multiple choice tests—whether they are the old paper and pencil versions or the new, slick computer versions—end up being taught in curricula that look like multiple choice tests. They spend a lot of time identifying and regurgitating information instead of creating and applying information in their talking with each other and their writing for each other.

If we want students to become sophisticated in their written and spoken explanations and arguments, as well as in their abilities to engage in evidence-based critique and debate, then we need to assess these sophisticated skills with performance assessments. Multiple choice testing will not get us there—no matter how slickly it is conceived for computer applications, including for adaptive testing in which students are given items scaled to be more or less difficult based on their answers. Multiple choice testing will keep us, our students, and our teachers in the same old identification and regurgitation ruts that NCLB institutionalized in state tests and consequently in instruction.

My next blog entry will be a deeper dive into the ways that we can incorporate performance assessments and portfolios in actual taught curriculum without creating havoc and expense.